Moving Object Tracking in Clinical Scenarios: Application to Cardiac Surgery and Cerebral Aneurysm Clipping

We are very much thankful to the reviewers for their thorough review. We have thoroughly revised the present manuscript in the light of their useful suggestions and comments. We hope the revision has improved the manuscript to a level of their satisfaction. Number wise answers to their specific comments/suggestions/queries are as follows (Table I, II, and III).

TABLE I
REVIEWER(2)'S COMMENTS

Sl. No.	Reviewer(2)'s Comments	Corresponding Response
1	Section 1.1: This section, which is supposed to motivate the presented method, present a host of different applications where real-time motion tracking on medical images supposedly may have a use. However, in most of the examples it is in no way obvious how motion tracking can help. These examples should be removed.	We thank the reviewer for their very constructive comments. As rightly pointed out by the reviewer, this section is supposed to motivate the presented method, by presenting a host of different surgical applications where real-time motion tracking on medical images could have a use. Accordingly, and based on further reviewer comments below, we have now added relevant and substantiated examples of use-cases for tool tracking/analysis within the realm of surgery.
2	P. 2, p. 2: I don't agree that tool manipulation and positioning "require augmented support", even though they may perhaps benefit from it.	We do agree with the reviewers comment that tool manipulation and positioning might not "require augmented support", but we do believe that it would be greatly beneficial. To clarify this further, we have directed the focus of the application to training and planning in minimally invasive surgery (MIS). In comparison to open surgery, MIS skills are more complex, nuanced and time-dependent to develop due to restricted vision, limited working space, loss of visual cues and tactile feedback. Thus MIS surgical training and planning could benefit greatly from visual support provided by instrument/motion tracking, by providing benchmarked metrics for continued objective and constructive assessment of highest standards of surgical skills, and lowering the risk of false tool trajectories and orientations [7], alignment of implants and placement of screws [8] etc.
3	P. 2, p. 3: The claim that "this support can be provided though object/motion tracking of the tools" is also unsubstantiated.	We apologize for not having substantiated our claims in the original draft. We have now clarified the focus of the method within surgery, and substantiated the claims of support by suggesting metrics (path length/deviation, economy and smoothness of hand movements, depth perception, rotational orientation, changes in instrument velocity and time) that can be used to provide this support in MIS training and planning.

TABLE II
REVIEWER(2)'S COMMENTS

Sl. No.	Reviewer(2)'s Comments	Corresponding Response
		We have also substantiated the claims by providing references to journal articles which propose that intra-operative tracking/detection of surgical instruments can provide important information to monitor instruments for the operational navigation in MIS leading to (a) ensured patient safety via proficient tool movements and avoidance of critical tissue structures, and (b) facilitation of a smooth and efficient invasive procedure [12].
4	P. 2, p. 3: It is stated that "the main goals of tool positioning are to (a) ensure safety of the patient, (b) avoid compression injuries and (c) facilitate a smooth and efficient surgery". However, it is not substantiated in any way that object tracking in itself can contribute to any of these.	We thank the reviewer for pointing this out. We believe that these are the perceived goals of tool tracking/positioning based on claims from published literature exploring applications of surgical navigation systems and tool tracking/motion analysis within surgery. Hence, we have now substantiated our text by providing references to supporting articles, as well as metrics that be used to objectively assess surgical training and planning which directly impact patient's intra-operative and postoperative outcome.
5	P. 2, p. 3: We can agree that "precise positioning of the tools remains pivotal in minimally invasive surgical procedures", but how this is "highlighting the need of object tracking" is left unsaid.	As mentioned in the response above and substantiated by the examples and references in the current version of the paragraph, precise positioning of the tools remains pivotal in minimally invasive surgical procedures [13] highlighting the need of object tracking via its impact on surgical training and intra-operative guidance.
6	P. 2, p. 4: Here, it is claimed that "In laparoscopic cholecystectomy (removal of the gall bladder), tool tracking can help the surgeon expertly maneuver his tools/instruments and avoid biliary and cystic duct injuries, bowel perforations, arterial bleeding and post operative infections and scarring [4]". However, this is again a highly dubious claim, and the reference [4], which gives the impression of supporting the claim, simply describes the frequency of such complications.	We thank the reviewer for pointing out this oversight from our end, and we apologize for the same. We have removed this example, and provided examples and references for new MIS use-cases in urology, orthopedic surgery, interventional radiology, ENT surgery, craniomaxillofacial surgery and cardiothoracic surgery.

TABLE III
REVIEWER(2)'S COMMENTS

Sl. No.	Reviewer(2)'s Comments	Corresponding Response
7	P. 2, p. 4: Just by reading the title of reference [6] it becomes clear that it does not "explore benefits of motion tracking in Otologic/ENT Surgery", but rather explores its use for "objective skills assessment in otologic surgery". You could ofcourse argue that objective skills assessment is indirectly beneficial to the surgery, but that is a strecht, and the claim as it stands here is at least highly misleading.	We thank the reviewer for pointing out this oversight from our end, and we apologize for the same. We have removed this example, and provided examples and references for new MIS use-cases in urology, orthopedic surgery, interventional radiology, ENT surgery, craniomaxillofacial surgery and cardiothoracic surgery. However, we disagree with the point that objective skills assessment is only indirectly beneficial to the surgery. Quality and adequacy of surgical proficiency directly impacts intraoperative and postoperative outcomes, and objective skills assessment plays a big part in assessing this proficiency [6].
8	Section 3.1: The results and the discussion is still mixed. The paper really should have a separate, short section describing the results and a separate section where the results are discussed.	We sincerely thank the reviewer for the suggestion. Now, we have included another "Discussion" subsection besides "Results" subsection.
9	Fig. 8: The (a) and (b) in the figure caption are wrong. If I understand it correctly, it should read "(a)Perceptual quality measures by Fourier (Method 2) and MODWT (Method 1). (b) Energy convergence comparison of three methods." Also, it the explanation of what the m on the x-axis is, that is now included in the text, should also be included here.	We are extremely sorry for the error; now we have fixed that in the revised manuscript.

Noname manuscript No.
(will be inserted by the editor)

Moving Object Tracking in Clinical Scenarios: Application to Cardiac Surgery and Cerebral Aneurysm Clipping

Sarada Prasad Dakua* · Julien Abinahed · Ayman Zakaria · Shidin Balakrishnan · Georges Younes · Nikhil Navkar · Abdulla Al-Ansari · Xiaojun Zhai · Faycal Bensaali · Abbes Amira

the date of receipt and acceptance should be inserted later

Abstract Background and Objectives
Surgical procedures, such as Laparoscopic and Robotic, are popular since they are invasive in nature and use miniaturized surgical instruments for small incisions. Tracking of the instruments (graspers, needle drivers) and field of view from the stereoscopic camera during surgery could further help the surgeons to remain focussed and reduce the probability of committing any mistakes. Tracking is usually preferred in computerized video surveillance, traffic monitoring, military surveillance system, and vehicle navigation. Despite the numerous efforts over last few past years, object tracking still remains an open research problem, mainly due to motion blur, image noise, lack of image texture, and occlusion. Most of the existing object tracking methods are time-consuming and less accurate when the input video contains high volume of information and more number of instruments.

Methods
This paper presents a variational framework to track the motion of moving objects in surgery videos. The key contributions are: 1- a denoising method using stochastic resonance in maximal overlap discrete wavelet transform is proposed, and 2- developing a robust Energy functional based on Bhattacharyya coefficient to match the target region in the first frame of the input sequence with the subsequent frames using a similarity metric. A modified affine transformation based registration is used to estimate the motion of the features following an active contour based segmentation method to converge the contour resulted from the registration process.

Results and Conclusion
The proposed method has been implemented on publicly available databases; the results are found satisfactory. Overlap index (OI) is used to evaluate the tracking performance and the maximum OI is found to be 76 % and 88 % on private data and public data sequences.

Department of Surgery, Hamad Medical Corporation
Doha, Qatar, E-mail: sdakua@hamad.qa

Department of Computer Science and Engineering, Qatar University
Doha, Qatar

School of Computer Science and Electronic Engineering, University of Essex
UK

Keywords Cerebral aneurysm · segmentation · object tracking · heart surgery · brain Aneurysm clipping · level sets

1 Introduction

Looking at the steep rise in cardiac diseases, bona fide treatment including surgery is necessary to prevent its rise and avoid sudden cardiac death [1]. Similarly, cerebral aneurysm (CA) is one of the devastating cerebrovascular diseases of adult population worldwide that causes subarachnoid hemorrhage, intra cerebral hematoma and other complications leading to high mortality rate [2]. Surgery is considered as an efficient modality for the patients with cardiac complications and ruptured cerebral aneurysms. Tracking could be considered as a treatment support and planning in Robotic, Laparoscopic and medical education. During Robotic surgery or Laparoscopic surgery, the surgeons concentrate on the surgery to avoid even slighted possible mortality and morbidity and usually get stressed. In this scenario, motion tracking of the tools and viewing the desired operating field may be considered two supportive pillars to augment the treatment and improve success rate.

1.1 Clinical Requirements in Surgery

Many factors contribute to successful outcome of a surgery, specifically minimally invasive surgery (MIS). These include technical factors, such as in-depth understanding of the relevant anatomy, clear understanding of the steps involved in the procedure, well-honed surgical skills and tool manipulation, as well as anthropomorphic factors such as operating team chemistry and dynamics. To a certain degree, MIS surgeons can advance their anatomy knowledge and procedural understanding through reading and surgical videos; however, other technical skills such as tool manipulation and positioning, which are very crucial to the successful outcome of the surgery [3], [4] are more complex, nuanced and time-dependent to develop due to restricted vision, limited working space, loss of visual cues and tactile feedback [5]. Quality and adequacy of surgical proficiency directly impacts intra-operative and postoperative outcomes [6]. The existing "apprenticeship" model of training in surgery provides limited and time-consuming opportunities to gain the required technical competencies. In its current form, the assessment of surgical proficiency is heavily reliant on subject-matter experts/subjective assessments [3]. Thus surgical training and planning could benefit greatly from visual support provided by instrument/motion tracking, by providing benchmarked metrics for continued objective and constructive assessment of highest standards of surgical skills, and lowering the risk of false tool trajectories and orientations [7], alignment of implants and placement of screws [8] etc.

Such augmented visual support for both surgical training and planning could be provided though object/motion tracking of the tools (such as scope, scissors, etc.) by providing objective assessment, benchmarking, and automated feedback on metrics such as path length/deviation, economy and smoothness of hand movements, depth perception, rotational orientation, changes in instrument velocity and time [9]. Zhao et al. [10] report that intra-operative tracking/detection of surgical instruments can provide important information to monitor instruments for the operational navigation in MIS, especially in the robotic minimally invasive surgeries (RMIS). Thus based on the above, the perceived impact of tool tracking/positioning on surgical training and intra-operative guidance leads to (a) ensured patient safety via proficient tool movements and avoidance of critical tissue structures, and (b) facilitation of a smooth and efficient invasive procedure [11]. This is crucial in surgery, as by continuously charting the location, movement, speed and acceleration of the different surgical instruments in the operating field, the surgeon

is continuously aware of the whereabouts of his instruments in relation to the patient's vital organs, blood vessels and nerves during surgery. For surgical training, it objectively helps assess surgical performance and helps differentiate between an expert and a novice surgeon, such that optimal training can then be provided to the novice to ensure the highest levels of patient care [3]. Therefore, precise positioning of the tools remains pivotal in minimally invasive surgical procedures [12] highlighting the need of object tracking via its impact on surgical training and intra-operative guidance.

Kobayashi et al. [13] applied surgical navigation techniques and tool tracking to renal artery dissection within the robot assisted partial nephrectomy procedure, and found that inefficient tool movements involving "insert", "pull", and "rotate" motions, as well as time to visualize and dissect the artery were significantly improved owing to improved visualization and control over the tool and anatomy. Pediatric orthopedic surgeons found an increase in accuracy and a reduction in operating time when using image guided surgical robotic systems to overcome the inaccuracies of hand-controlled tool positioning [14], these robots achieve this by providing information about surgical tools or implants relative to a target organ (bone). In Urology, motion tracking can greatly assist in outpatient procedures such as MRI and ultrasound guided prostate biopsy, allowing the surgeon to accurately position and invade suspicious malignant zones for a tissue sample [16]. In interventional radiology, motion tracking can help track guide-wires during endovascular interventions and radiation therapy [17]. In addition to these, applications of surgical navigation systems and tool tracking/motion analysis are being explored in many other surgical fields, including ear-nose-and-throat (ENT) surgery [7], craniomaxillofacial surgery [18], cardiothoracic surgery [19], and orthopedic surgery [20].

1.2 Related Work

The literature of motion tracking is rich; a few recent methods are included in this paper. Kim and Park [21] present a strategy that is based on edge information to assist object-based video coding, motion estimation, and motion compensation for MPEG 4 and MPEG 7 utilizing the human visual perception to provide edge information. However, the method critically depends on its ability to establish correct correspondences between points on the model edges and edge pixels in an image. Furthermore, this is a non trivial problem especially in the presence of large inter frame motions and cluttered environments. Subudhi et al. [22] propose a two step method: spatio-temporal spatial segmentation and temporal segmentation that uses Markov random field (MRF) model and posteriori probability (MAP) estimation technique. Duffner and Garcia [23] present an algorithm for real time single object tracking, where a detector makes use of the generalized Hough transform with color and gradient descriptors; a probabilistic segmentation method is used for foreground and background color distributions. However, it is computationally expensive, especially, when the number of parameters is large. It also could be erroneous because the gradient information usually leads to error when noise level is high. Li et al. [24] suggest a method within the correlation framework (CF) that models a tracker maximizing the margin between the target and surrounding background by exploiting background information effectively. They propose to train a CF by multilevel scale supervision, which aims to make CF sensitive to the target scale variation. Then the two individual modules are integrated into one framework simplifying the tracking model. However, the computational load and efficiency are still two major concerns. Mahalingam et al. [25] propose a fuzzy morphological filter and blob detection based method for object tracking. However, the performance gets deteriorated in presence of noise, lack of illumination and occlusion. Zhang et al. [26] propose a correlation particle filter (CPF) that combines a correlation filter and a particle filter. However, this tracker is still unable to deal

with scale variation and partial occlusion. Yang et al. [27] present a method to analyze frames extracted from videos using Kernelized Correlation Filters (KCF) and Background Subtraction (BS) (KCF-BS) to plot the 3D trajectory of cabbage butterfly. The KCF-BS algorithm is used to track the butterfly in video frames and obtain coordinates of the target centroid in two videos. However, it is noticed that the target sometimes gets lost and the method is unable to re-detect or recognize the target when the target motion is fast. Du et al. [28] propose an object tracking method for satellite videos by fusing KCF tracker and a three frame difference algorithm. Although the method reports interesting results, it takes long time to perform. Liu et al. [30] propose a correlation filter based tracker that consists of multiple positions' detections and alternate templates. The detection position is repositioned according to the estimated speed of target by an optical flow method and the alternate template is stored with a template update mechanism. However, this method fails to perform if the size of each target is too small compared with the entire image, and the target and the background are very similar. Liu et al. [31] propose a method by integrating histogram of oriented gradient, RGB histogram and motion histogram into a novel statistical model to track the target in unmanned aerial vehicle captured videos. However, it fails to perform in occluded scenes.

Du et al. [32] present a method that is based on iterative graph seeking. Usually the superpixel based methods use mid-level visual cues to represent target parts where local appearance variations are exploited by superpixel representation. These based methods have three sequential steps: (A) target part selection, (B) target part matching, and (C) target state estimation. (A) selects candidate target parts from the background, (B) a local appearance model associates parts between consecutive frames (target part matching), center pixel location and size of the target) is estimated based on majority voting, (C) target state is estimated based on majority voting of matching results. This method integrates target part selection, part matching, and state estimation using a unified energy minimization framework. It incorporates structural information in local parts variations using the global constraint. Although the results are reported promising, the target part selection and target part matching when combinedly merge with the correlation filter, the estimation of the target takes long time to converge due to scale variation and partial occlusion that are bound to happen in surgery scenarios. Furthermore, when the noise level (for instance, in cardiac cine MRI data) in the input frames is high, the method would certainly struggle to perform. We intend to address these issues through our proposed method. Furthermore, if the literature above is carefully observed, noise has always been an issue in most of the methods. Therefore, in our proposed method, we first denoise the input frames. The target region on the first frame is chosen by a level set (LS) function and then the foreground & background models are generated. The foreground and background distributions are determined using the models in subsequent frames and the motion of the pixels from the region of interest are estimated through a registration framework. Additionally, the selected region contour in the current frame is registered with the subsequent frame. Finally, segmentation is applied to refine the contour generated during registration and the contour is updated.

The paper is organized as follows: Section 2 describes the denoising stage, Section 2.2 presents the approach for target rendering (including region selection and developing models), Section 2.3 defines a method for motion estimation through registration, Section 2.4 presents the segmentation, Section 3 provides the results while Section 4 concludes the paper.

2 Methodology and Data

The method is illustrated in Fig. 1. First, the input frame is denoised to minimize the the negative impact of noise on subsequent steps. The target region is then selected following by the devel-

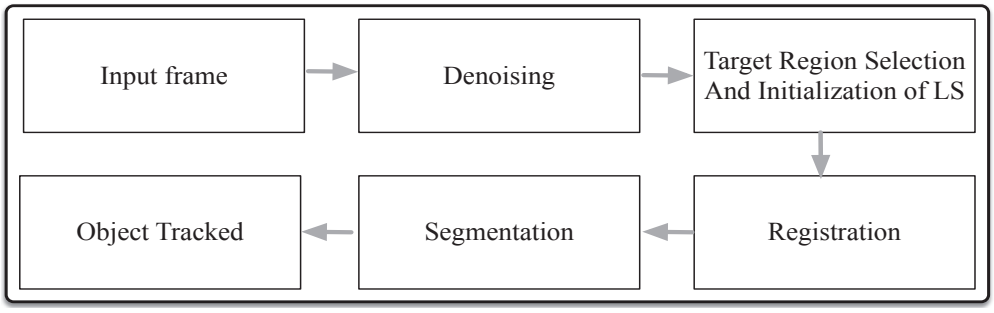


Fig. 1 Block Diagram describing the proposed method.

opment of background models for motion estimation through a registration framework. Finally, the rough contour generated in registration step is further refined (by a proper segmentation method) and the contour is updated on subsequent frames.

2.1 Denoising of Image Sequences

Over the years, most of the methods address the noisy and cluttered medical images, mostly, by filtering that result significant degradation in image quality. One of the efficient approaches that counter noise and constructively utilize noise is Stochastic Resonance (SR) [34]. SR occurs if the Signal to Noise Ratio (SNR) and input/output correlation have a well marked maximum at a certain noise level. Unlike very low or high noise intensities, moderate ones allow the signal to cross the threshold giving maximum SNR at some optimum noise level. In the bistable SR model, upon addition of zero mean Gaussian noise, the pixel is transferred from weak signal state to strong signal state, which is modeled by Brownian motion of a particle (*pc*) placed in a double well potential system. The state at which performance metrics are found optimum can be considered as the stable state providing maximum SNR. There have already been many attempts to use SR in different domains such Fourier and spatial domains [42]; however, we have chosen the maximal overlap discrete wavelet transform (MODWT) [37] because of some of its key advantages: 1) MODWT can handle any sample size, 2) the smooth and detail coefficients of MODWT multiresolution analysis are associated with zero phase filters, 3) it is transform invariant, and 4) it produces a more asymptotically efficient wavelet variance estimator than DWT.

2.1.1 Maximal Overlap Discrete Wavelet Transform

Generally, DWT is defined by: $\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi\left(2^j t - k\right) \quad j,k \in \mathrm{Z};\ z = \{0,1,2...\}$, where ψ is a real valued function compactly supported, and $\int_{-\infty}^{\infty} \psi(t)\,dt = 0$. MODWT is evaluated using dilation equations: $\phi(t) = \sqrt{2}\sum_k l_k \phi(2t-k)$, $\psi(t) = \sqrt{2}\sum_k h_k \phi(2t-k)$, where $\phi(2t-k)$ and $\psi(t)$ are father wavelet defining low pass filter coefficients and mother wavelet defining high pass filter coefficients l_k: $l_k = \sqrt{2}\int_{-\infty}^{\infty} \phi(t)\,\phi(2t-k)\,dt$, $h_k = \sqrt{2}\int_{-\infty}^{\infty} \psi(t)\,\psi(2t-k)\,dt$.

2.1.2 Denoising by MODWT

In this methodology, 2D MODWT is applied to the $M \times N$ size image I. Applying SR to the approximation and detail coefficients, the stochastically enhanced (tuned) coefficient sets in MODWT domain are obtained as $W_{\psi}^{s}(l,p,q)_{SR}$ and $W(l_0,p,q)_{SR}$. The SR in discrete form is defined as: $\frac{dx}{dt} = \left[ax - bx^3\right] + B\sin\omega t + \sqrt{D}\xi(t)$, where $\sqrt{D}\xi(t)$ and $B\sin\omega t$ represent noise and input, respectively; these are replaced by MODWT sub-band coefficients. The noise term is the factor to produce SR; maximization of SNR occurs at the double well parameter a. Implementation of SR on digital images necessitates the need for solving the stochastic differential equation using Euler-Maruyama's method [36] that gives the iterative discrete equation:

$$x(n+1) = x(n) + \Delta t\left[(ax(n) - bx^3(n)) + Input(n)\right] \tag{1}$$

where a and b are the bistable parameters, whereas n and Δt represent iteration and sampling time, respectively. $Input$ denotes the sequence of input signal and noise, with the initial condition being $x(0) = 0$. $Input$ denotes the sequence of input signal and noise, with the initial condition being $x(0) = 0$. The final stochastic simulation is obtained after some predefined number of iterations. Given the tuned (enhanced and stabilized) set of wavelet coefficients ($X_{\phi}(l_0,p,q)$ and $X_{\psi}^{s}(l,p,q)$), the denoised image $I_{denoised}$ in spatial domain is obtained by inverse maximal overlap discrete wavelet transform (IMODWT) as:

$$I_{denoised} = \frac{1}{\sqrt{MN}}\sum_{p}\sum_{q} X_{\phi}(l_0,p,q)\,\phi_{l_0,p,q}(i,j) + \frac{1}{\sqrt{MN}}\sum_{s\in(H,V,D)}\sum_{l=l_0}\sum_{p}\sum_{q} X_{\psi}^{s}(l,p,q)\,\psi_{l_0,p,q}^{s}(i,j)$$

The double well parameters a and b are determined from the SNR by differentiating SR with respect to a and equating to zero; in this way, SNR is maximized resulting in $a = 2\sigma_0^2$ for maximum SNR, where σ_0 is the noise level administered to the input image. The maximum possible value of restoring force ($R = B\sin\omega t$) in terms of gradient of some bistable potential function $U(x)$, $R = -\frac{dU}{dx} = -ax + bx^3$, $\frac{dR}{dx} = -a + 3bx^2 = 0$ resulting $x = \sqrt{a/3b}$. R at this value gives maximum force as $\sqrt{\frac{4a^3}{27b}}$ and $B\sin\omega t < \sqrt{\frac{4a^3}{27b}}$. Maximizing the left term (keeping $B = 1$), $b < \frac{4a^3}{27}$. In order to get the parameter values, $a = w \times 2\sigma_0^2$, and $b = z \times \sqrt{\frac{4a^3}{27}}$; w and z are weight parameters for a and b, initially, w is an experimentally chosen constant that later becomes input image standard deviation dependent, while z is a number less than 1 to ensure sub-threshold condition of the signal. In this way, the noise in input image is countered and maximum information from the image is achieved.

2.2 Target Rendering

Target region selection or target rendering [38] [29] is the initial step in this motion tracking. Then the features (such as intensity, color, edge, texture, etc.) are selected that can appropriately describe the target. The notations used in target rendering are: fs- feature space, r- number of features, foreground distribution (by the features)- **fd**, and background distribution- **bd**. The region is initialized on the first frame and represented by a level set function ϕ because of its flexibility in choosing the contour. The distributions of foreground ($\phi \geq 0$) and background ($-th < \phi < 0$, th is the threshold to restrict the region of interest into small area) regions are represented by $fg(\phi)$ and $bg(\phi)$, respectively and match with **fd** and **bd**. Next, the foreground and background models are generated. Suppose the pixels $\{x_{f,i}\}_{i=1,\ldots,n_f}$ and $\{x_{b,i}\}_{i=1,\ldots,n_b}$ fall in foreground and background regions, the function $z : \Re^2 \rightarrow \{1,\ldots,r\}$ can be used to map the pixels (x_i) into the bin $b(x_i)$ in feature space. The probability of the feature space in the models

are: $fd_{fs} = \frac{1}{n_f} \sum_{i=1}^{n_f} \delta\left[(x_{i,f}) - fs\right]$ and $bd_{fs} = \frac{1}{n_b} \sum_{i=1}^{n_b} \delta\left[(x_{i,f}) - fs\right]$, where δ is the Kronecker delta function and n_f and n_b are the number of pixels in foreground and back ground, respectively. The foreground and background distributions in the current frame candidate region ($-th < \phi < 0$) are obtained as:

$$fg\left(\phi\right) = \frac{1}{F_f} \sum_{i=1}^{n} H\left(\phi\left(x_i\right)\right) \delta\left[b\left(x_i\right) - fs\right] \; and \; bg\left(\phi\right) = \frac{1}{F_b} \sum_{i=1}^{n} \left(1 - H\left(\phi\left(x_i\right)\right)\right) \delta\left[b\left(x_i\right) - fs\right] \quad (2)$$

$H(.)$ is the Heaviside function to select foreground region; F_f and F_b are the normalization factors.

2.3 Registration

Registration of the target in the first frame with the next subsequent frame is performed to estimate the affine deformation of the target. We determine the foreground and background distribution in the frames and match them with respective foreground and background models. We use Bhattacharyya metric [39] because it is computationally fast and is already being used in face recognitions for years. Additionally, it has straightforward geometric interpretation. Since it is the cosine angle between **fd** and $fg(\phi)$ or between **bd** and $bg(\phi)$, higher value of the coefficient indicates better matching between candidate and target models. Thus our similarity distance measure:

$$En_1\left(\phi\right) = \sum_{fs=1}^{r} \left(\sqrt{fg_{fs}\left(\phi\right) fd_{fs}} + \gamma \sqrt{bg_{fs}\left(\phi\right) bd_{fs}}\right) \quad (3)$$

where γ is the weight to balance the contribution from both foreground and background in the matching.

For deformation estimation, we have proposed a simple and efficient framework as follows. Suppose in the current frame, ϕ_0 is the target initial position, the contour is obtained by $\phi = 0$. The probabilities $fg\left(\phi_0\right) = \left\{fg_{fs}\left(\phi_0\right)\right\}_{fs=1,\ldots,r}$ and $bg\left(\phi_0\right) = \left\{bg_{fs}\left(\phi_0\right)\right\}_{fs=1,\ldots,r}$ are computed. Applying Taylor's expansion:

$$\begin{aligned} En_1\left(\phi\right) = \tfrac{1}{2} \left(\sum_{fs=1}^{r} \sqrt{fg_{fs}\left(\phi_0\right) fd_{fs}} + \sum_{fs=1}^{r} fg_{fs}\left(\phi\right) \sqrt{\tfrac{fd_{fs}}{fg_{fs}(\phi_0)}}\right) \\ + \tfrac{1}{2}\gamma \left(\sum_{fs=1}^{r} \sqrt{bg_{fs}\left(\phi_0\right) bd_{fs}} + \sum_{fs=1}^{r} bg_{fs}\left(\phi\right) \sqrt{\tfrac{bd_{fs}}{bg_{fs}(\phi_0)}}\right) \end{aligned} \quad (4)$$

By putting equation (2) in (4), we get:

$$\begin{aligned} En_1\left(\phi\right) = \tfrac{1}{2} \left(\sum_{fs=1}^{r} \sqrt{fg_{fs}\left(\phi_0\right) fd_{fs}} + \tfrac{1}{F_f} \sum_{fs=1}^{n} h_{f,i} H\left(\phi\left(x_i\right)\right)\right) \\ + \tfrac{1}{2}\gamma \left(\sum_{fs=1}^{r} \sqrt{bg_{fs}\left(\phi_0\right) bd_{fs}} + \tfrac{1}{F_b} \sum_{fs=1}^{n} h_{b,i} \left(1 - H\left(\phi\left(x_i\right)\right)\right)\right) \end{aligned} \quad (5)$$

where the weights that play pivotal role in detecting the new centroid of the target are: $h_{f,i} = \sum_{fs=1}^{r} \sqrt{\frac{fd_{fs}}{fg_{fs}(\phi_0)}} \delta\left[z\left(x_i\right) - fs\right]$ *and* $h_{f,i} = \sum_{fs=1}^{r} \sqrt{\frac{bd_{fs}}{bg_{fs}(\phi_0)}} \delta\left[z\left(x_i\right) - fs\right]$. Higher value of Bhattacharyya coefficient can be obtained by maximizing (5) that is a function of location x and contour.

Furthermore, we consider the foreground and background intensity as additional feature. Suppose the first frame, $u_0(x,y)$, consists of two concentric regions (u_0^i, u_0^o) meaning the input image contain more than one intensity labels. This is certainly challenging in determining a smooth contour initialization & deformation because of varying intensities. Therefore, we integrate both local and global image information in the energy term in order to make it perform as a perfect step detector with respect to the initialization of contour. The energy term is defined as:

$$En_2 = \lambda_1 E^G + \lambda_2 E^L + E^R \tag{6}$$

where λ_1 and λ_2 are fixed constants; E^G, E^L, and E^R are the global term, local term, and regularized term, respectively (containing respective image information). E^R controls the boundary smoothness. The local term is defined as,

$$E^L = \int\limits_{\phi<0} \frac{(g_k u_0(x,y) - u_0(x,y) - d_1(x,y))^2}{d_1(x,y)^2} dxdy + \int\limits_{\phi>0} \frac{(g_k u_0(x,y) - u_0(x,y) - d_2(x,y))^2}{d_2(x,y)^2} dxdy \tag{7}$$

where g_k is an averaging filter with $k \times k$ size, d_1 and d_2 are intensity averages of the difference image $g_k u_0(x,y) - u_0(x,y)$ inside and outside the variable curve C, respectively. The global term:

$$E^G = \int\limits_{\phi<0} \frac{(u_0(x,y) - c_1(x,y))^2}{c_1(x,y)^2} dxdy + \int\limits_{\phi>0} \frac{(u_0(x,y) - c_2(x,y))^2}{c_2(x,y)^2} dxdy \tag{8}$$

where the constants c_1, c_2 represent as the average intensity of $u_0(x,y)$ inside C and outside C, respectively. c_1 and c_2 are approximated by a weighted average of image intensity $u_0(p,q)$, where (p,q) is the neighborhood of (x,y). It means $c_1(x,y)$ and $c_2(x,y)$ are spatially varying; we formulate $c_1(x,y)$ and $c_2(x,y)$ as, $c_1(x,y) = \frac{\int_\Omega g_k((x,y)-(p,q)) u_0(p,q) H(\phi(p,q)) dpdq}{g_k((x,y)-(p,q)) H(\phi(p,q)) dpdq}$ and $c_2(x,y) = \frac{\int_\Omega g_k((x,y)-(p,q)) u_0(p,q)(1-H(\phi(p,q))) dpdq}{g_k((x,y)-(p,q))(1-H(\phi(p,q)))dpdq}$. We use the conventional regularizing term E_R that includes a penalty term on the total length of the edge contour for a given segmentation. Also it contains another penalty term on the total area of the foreground region found by the segmentation. The energy term therefore becomes:

$$\begin{aligned} En_2(\phi) = \mu \int\limits_\Omega \delta(\phi) + v \int\limits_\Omega H(\phi(x,y)) dxdy + |\nabla\phi| dxdy + \lambda_1 \int\limits_\Omega \frac{(u_0(x,y) - c_1(x,y))^2 H(\phi(x,y))}{c_1(x,y)^2} dxdy + \\ \lambda_1 \int\limits_\Omega \frac{(u_0(x,y) - c_1(x,y))^2 (1 - H(\phi(x,y)))}{c_2(x,y)^2} dxdy + \lambda_2 \frac{(g_k u_0(x,y) - d_1(x,y))^2 H(\phi(x,y))}{d_1(x,y)^2} dxdy + \\ + \lambda_2 \frac{(g_k u_0(x,y) - d_2(x,y))^2 (1 - H(\phi(x,y)))}{d_2(x,y)^2} dxdy \end{aligned} \tag{9}$$

This equation (9) has to be maximized to obtain higher Bhattacharyya coefficient. The similarity distance measure now becomes:

$$En(\phi) = En_1(\phi) + En_2(\phi) \tag{10}$$

We model the motion of the target as affine transformation by introducing a wrap in (10):

$$x = h(x, \Delta T) = \begin{pmatrix} 1 + fg_1 & fg_3 & fg_5 \\ fg_2 & 1 + fg_4 & fg_6 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{11}$$

The column vector characterizes the change in poses. Substituting (11) in (10) and omitting the terms that are not a function of ΔT-incremental warp (represented ϕ), we obtain:

$$En(\phi) = \frac{1}{2F_f} \sum_{i=1}^{n} H(\phi(h(x, \Delta T))) w_{f,i} + \frac{1}{2F_b} \gamma \sum_{i=1}^{n} (1 - H(\phi(h(x, \Delta T)))) w_{b,i} \tag{12}$$

ΔT tends to 0, the estimation gets converged. In this way, the registration step iteratively estimates the shape change until it gets converged.

2.4 Segmentation

Since the tracker in the registration stage is still not able to extract the target contour properly, the registration result needs to be refined through segmentation. In order to do this, we optimize ϕ in equation (10) because the equation is a function of ϕ; in other words, $\frac{\partial En(\phi(x_i))}{\partial \phi(x_i)} = 0$. This is solved by well known steepest-ascent method: $\frac{\partial En(\phi(x_i),t)}{\partial t} = \frac{\partial En(\phi(x_i))}{\partial \phi(x_i)}$. We obtain:

$$\begin{array}{l} \frac{\delta\phi(x,y,t)}{\delta t} = \delta_{\in}(\phi)\left[\mu\nabla.\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - v + \lambda_1\left(\frac{(u_0(x,y)-c_2(x,y))^2}{c_2(x,y)^2} - \frac{(u_0(x,y)-c_1(x,y))^2}{c_1(x,y)^2}\right)\right] + \\ \lambda_2\left(\frac{(g_k u_0(x,y)-d_2(x,y))^2}{d_2(x,y)^2} - \frac{(g_k u_0(x,y)-d_1(x,y))^2}{d_1(x,y)^2}\right) + \frac{1}{2}\Delta t\delta_{\in}(\phi)\left(\frac{1}{F_f}h_{f,i} - \gamma\frac{1}{F_b}h_{b,i}\right) \end{array} \tag{13}$$

$$\frac{\partial_{\varepsilon}(\phi)}{|\nabla\phi|}\frac{\partial\phi}{\partial\overrightarrow{n}} = 0 \; on \; \partial\Omega \tag{14}$$

where H and $\delta_{\in}$ represent the Heaviside function and Dirac measure, respectively; $\frac{\partial\phi}{\partial\overrightarrow{n}}$ and $\overrightarrow{n}$ denote the normal derivative of ϕ at the boundary and the exterior normal to the boundary, respectively. Finally, the target is updated on subsequent frames.

2.5 Data

The datasets used in this work are obtained from private sources such as Hamad Medical Corporation (30 data sequences) and public sources such as Sunnybrook [33] (45 data sequences) and VOT 2015 [41] (60 data sequences). The Sunnybrook Cardiac Data (SCD) consist of cine MRI data from a mixed of patients and pathologies: healthy, hypertrophy, heart failure with infarction and heart failure without infarction. Subset of this data set was first used in the automated myocardium segmentation challenge from short-axis MRI. The VOT 2015 sequences are chosen from a large pool of sequences including ALOV, OTB, non-tracking, Computer Vision Online, Professor Bob Fisher's Image Database, Videezy, Center for Research in Computer Vision, University of Central Florida, USA, NYU Center for Genomics and Systems Biology, Data Wrangling, Open Access Directory and Learning and Recognition in Vision Group, INRIA. The initial pool of sequences is created by combining the sequences from all the sources. After removal of duplicate sequences, grayscale sequences and sequences that contained objects with area smaller than 400 pixels, the final sequences are obtained; more details can be obtained from the website [47].

3 Results and Discussion

3.1 Results

The proposed method is implemented on both private and public databases as described earlier. The qualitative results of denoising are provided in Fig. 2. We have quantitatively compared the proposed denoising method with that of Fourier because of its huge popularity [42]. The perceptual quality measurement (PQM) [43] are provided in Fig. 8, which shows greater value in case of MODWT suggesting higher efficacy of MODWT; in this figure, m denotes mass of the particle that moves under stochastic condition. For denoisng of the input images, the initial values of Δt and z are taken as 0.007 and 0.000027, respectively. To determine the quality of the

denoised image, we have calculated distribution separation measure that estimates the degree of image quality. The DSM is defined as [42]: $DSM = \left|\mu_T^E - \mu_B^E\right| - \left|\mu_T^O - \mu_B^O\right|$, where μ_T^E and μ_T^O are the mean of the selected target regions of the denoised and original images, respectively; μ_B^E and μ_B^O are the mean of the selected background region of the denoised and original image, respectively. Higher the value of DSM, better is the quality. It is observed that the value of DSM is maximum at iteration 200 and then it starts decreasing, therefore, this iteration is considered as the optimal.

These denoised frames are further used in the subsequent steps in the proposed method. As mentioned earlier, we have included the image sequences of cardiac surgery and clipping for ruptured cerebral aneurysms in this work. We have also tested our method on cardiac cine MRI datasets, high contrast and low contrast level to highlight the performing capability of the method in varying intensities. The performance results on these datasets are provided in Fig. 5 and 7. We have chosen different scenarios for cerebral aneurysm surgical procedure (clipping): one is to track the scissors' or clippers' movement and the other one is to focus on the operating field during surgery, where multiple tools are used by the surgeons. It is important to track the motion of the scissors in order to minimize the damage caused by their movement. Besides the tools' tracking, capturing or tracking the operating field is also important; it helps the surgeon in concentrating on the tools used during the surgery and the impacted tissues of interest. The results are given in Fig. 3 and 4. We have also tested the proposed method on VOT 2015 datasets and found some satisfactory results as can be observed in Figure 6. We have included this particular dataset in this paper to emphasize on the fact that the foreground is not very significantly different than the background like it happens in medical data sequences. Usually, the medical data are blurry (either redish or grayish) and lack contrast as can be observed from the figures. In this scenario, only a contour surrounding the tools could easily be ignored, therefore, just for user's (surgeon) convenience, we have added the blue line surrounding the red line in the tracking results. While calculating the accuracy, red line is only taken into consideration. In order to determine the segmentation accuracy, we have used Dice coefficient (DC), which may be defined as [46]: $DC = 2 \times \frac{|X \cap Y|}{|X|+|Y|}$, where X and Y are two point sets. The average segmentation accuracy on 3 T machine is 94 % whereas in case of 7 T, it is found to be 96 %. The proposed method has performed as expected, which can be verified from the results provided in Section 3.1. We have optimized the algorithm and code; average time taken to perform tracking and average number of frames are less than 25-30 seconds and 24 frames per second, respectively. We have also compared the performance of the proposed method with other similar methods ([44],[32]); the proposed method converges faster than the other methods 8(b). We have also calculated overlap index (OI) [45] to determine the overlap between the resulted target contour and the actual boundary. We have found it highest in case of the proposed method against others as can been observed from Table 1.

3.2 Discussion

The values of bistable system parameters play a crucial role in the process of denoising using SR. The expression for SR on any data set contains additive terms of multiples of w and subtractive term of multiples of z. This is observed that the images that have low contrast and low dynamic range require larger values of w while those that have relatively more contrast and cover an appreciable gray level range require smaller values of w for proper denoising. Values of Δt have been studied to be similar to that of w. This is also perceived that w is inversely proportional to overall variance signifying the contrast of input image. Optimization process leads us to the optimum value of w; the value of z should be less than 1 so that condition $b < \sqrt{\frac{4a^3}{27}}$ holds

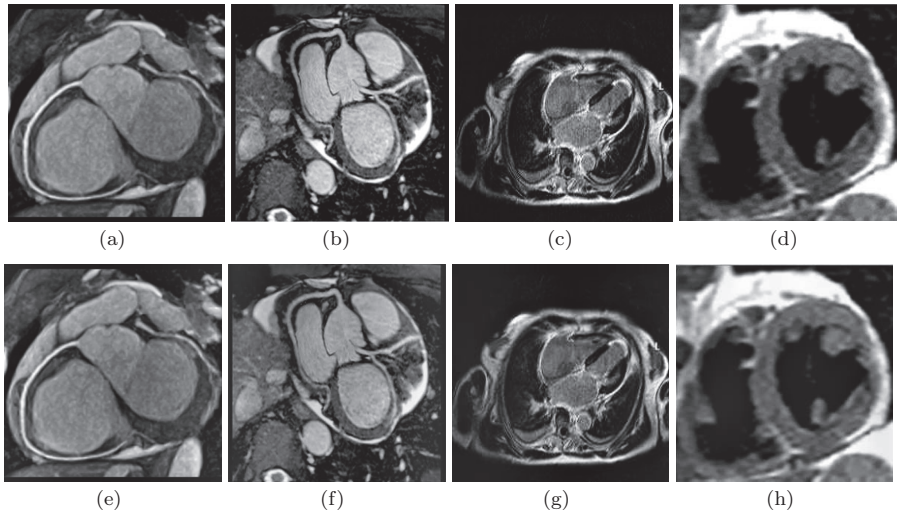


Fig. 2 (a-d) Input frames in a video sequence to be denoised. (e-f) Results of denoising.

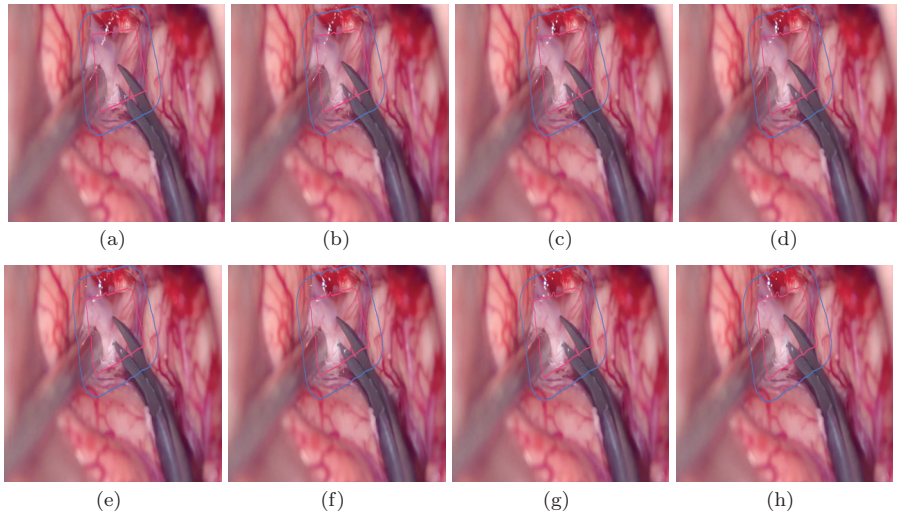


Fig. 3 Tracking of the operating field with multiple objects during cerebral aneurysm clipping.

assuring that the system is bistable and signal is sub-threshold so that SR can be applicable. We prefer a very small value of this factor to remain well within the allowable range of b. Finally, we have noticed that the varying segmentation accuracy depending on the quality of the input data sequence. The MRI data obtained from 7 T machine gives better accuracy than 3 T MRI machine.

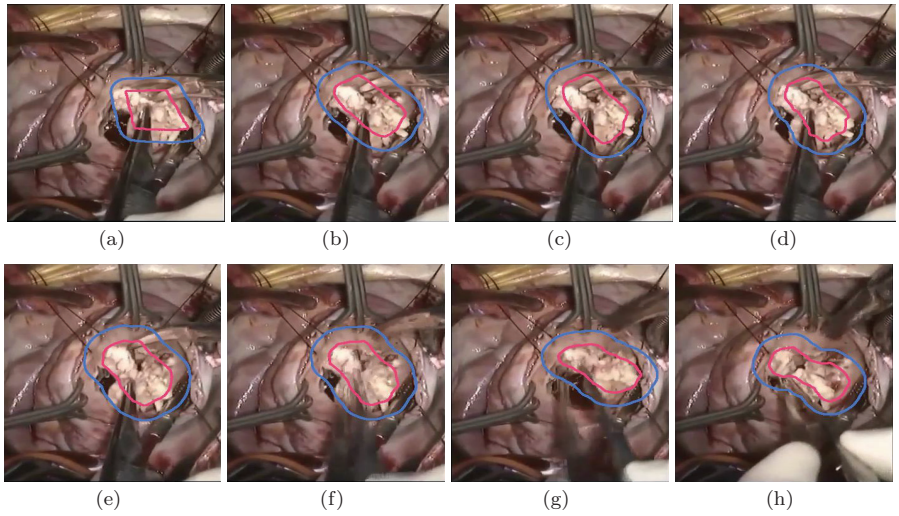


Fig. 4 Tracking of the operating field with multiple objects during cardiac surgery.

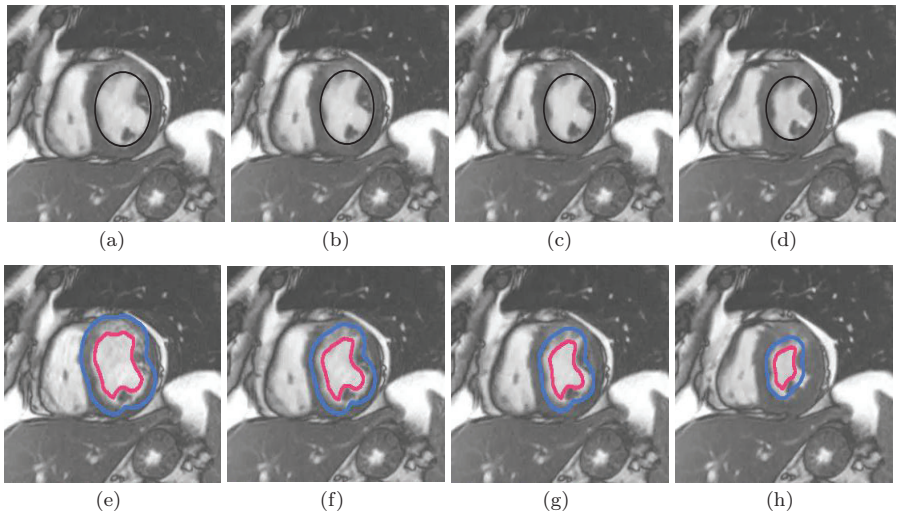


Fig. 5 (a-d) Ground truth frames. (e-h) Tracking of left ventricle in low-contrast cine magnetic resonance imaging (Low contrast-CMRI) during cardiac surgery.

4 Conclusions and Future Work

A variational framework has been presented to track the motion of moving objects and field of view in surgery sequences. We have presented a method that has used SR to denoise the input frames and a combined registration-segmentation framework to conduct motion tracking. We have introduced a robust similarity metric and an efficient energy functional in this framework. Despite the fact that the input data contain varying illumination, motion blur, lack of image texture, occlusion, and fast object movements, the performance of the proposed method is found

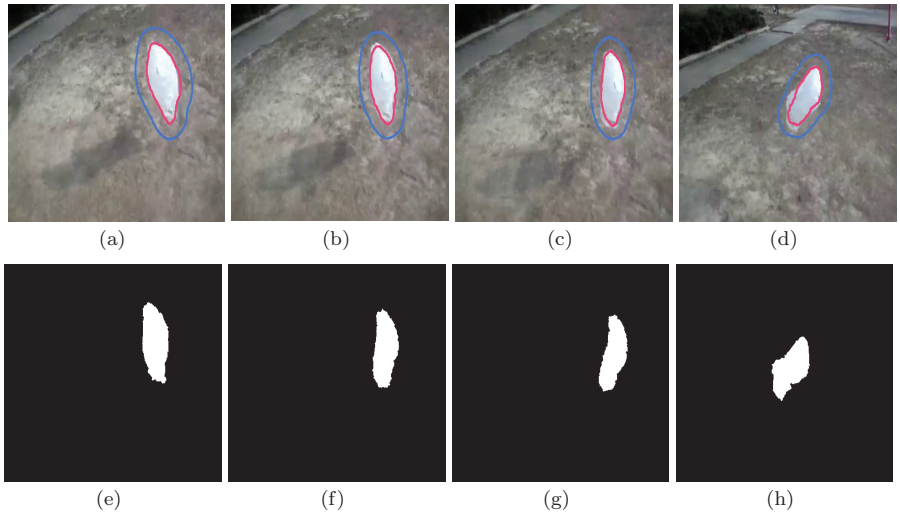


Fig. 6 (a-d) Tracked frames in a video sequence (VOT 2015). (e-f) Corresponding ground-truth sequences.

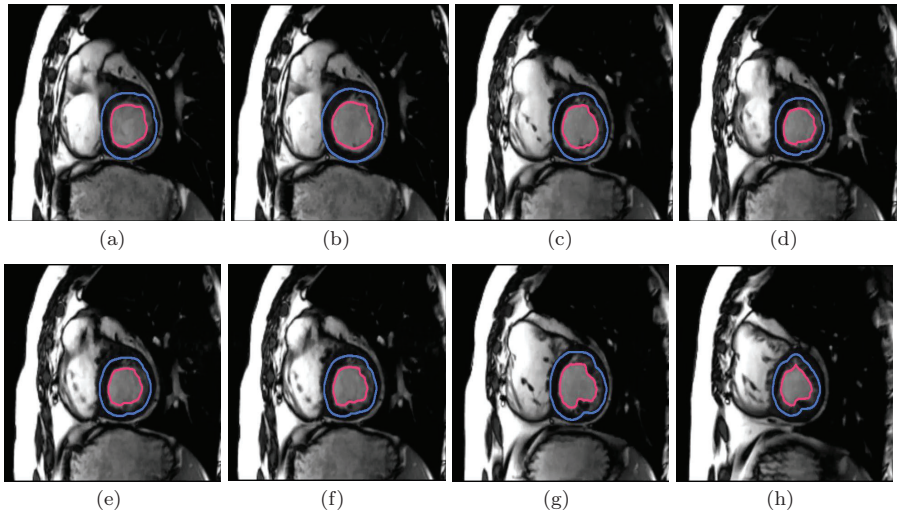
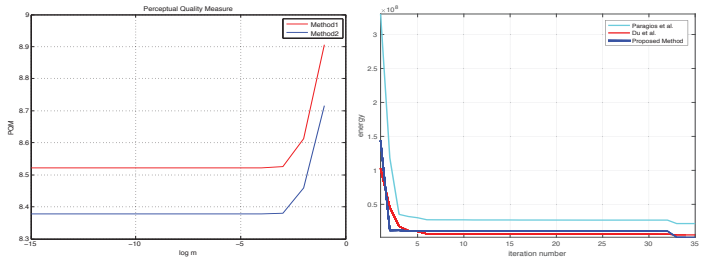


Fig. 7 Tracking of left ventricle in high contrast cine magnetic resonance imaging (High contrast-CMRI) during cardiac surgery.

quite satisfactory. In future, we intend to extensively evaluate the method quantitatively so that it can be well tested before trying in clinical practice.

Acknowledgement

This work was partly supported by NPRP Grant #NPRP 5-792-2-328 from the Qatar National Research Fund (a member of Qatar Foundation).



(a) (b)

Fig. 8 Perceptual quality measures by Fourier (Method 2) and MODWT (Method 1); m in x-axis denotes mass of the particle that moves under stochastic condition. (b) Energy convergence comparison of three methods.

Table 1 Overlap index comparison of different methods on hospital and VOT 2015 datasets.

	Hospital and SCD Dataset			
Method	FOV-CA	Scissors-CA	Low contrast-CMRI	High contrast-CMRI
Paragois et al. [44]	64 %	65 %	64 %	65 %
Du et al. [32]	67 %	69 %	68 %	72 %
Proposed Method	**72 %**	**74 %**	**73 %**	**76 %**
	VOT 2015 Dataset			
Method	Rabit	Shaking	Racing	Octopus
Paragois et al. [44]	69 %	70 %	68 %	71 %
Du et al. [32]	72 %	75 %	70 %	76 %
Proposed Method	**83 %**	**85 %**	**82 %**	**88 %**

References

1. T. G. Di Salvo, M. A. Acker, G. W. Dec, and J. G. Byrne, Mitral valve surgery in advanced heart failure, Journal of the American College of Cardiology, Volume 55, page 271-282 (2010).
2. X. Zhai, M. Eslami, E. S. Hussein, M. S. Filali, S. T. Shalaby, A. Amira, F. Bensaali, S. Dakua, J. Abinahed, A. Al-Ansari, and A. Z. Ahmed, Real-time automated image segmentation technique for cerebral aneurysm on reconfigurable system on chip, Journ. of Computational Science, Elsevier, Volume 27, page 35-45 (2018).
3. S. Ganni, S. M. B. I. Botden, M. Chmarra, A software-based tool for video motion tracking in the surgical skills assessment landscape, Surg Endosc, Volume 32, page 2994 (2018).
4. J. J. Jakimowicz and S. Buzink, Training Curriculum in Minimal Access Surgery, in Training in Minimal Access Surgery, N. Francis, A. Fingerhut, R. Bergamaschi, and R. Motson Eds. London: Springer London, page 15-34 (2015).
5. C. Feng et al., Surgical training and performance assessment using a motion tracking system, International Mediterranean Modelling Multiconference, I3M 2006, 2019, page 647-652 (2019).
6. S. M. Carroll, A. M. Kennedy, O. Traynor, and A. G. Gallagher, Objective assessment of surgical performance and its impact on a national selection programme of candidates for higher surgical training in plastic surgery, Journal of Plastic, Reconstructive and Aesthetic Surgery, Volume 62, page 1543-1549 (2009).
7. V. Pruliere-Escabasse and A. Coste, Image-guided sinus surgery, Eur Ann Otorhinolaryngol Head Neck Dis, Volume 127, page 33-39 (2010).
8. T. Tjardes et al., Image-guided spine surgery: state of the art and future directions, Eur Spine J, Volume 19, page 25-45 (2010).
9. S. Shaharan, E. Nugent, D. M. Ryan, O. Traynor, P. Neary, and D. Buckley, Basic surgical skill retention: Can patriot motion tracking system provide an objective measurement for it?, J Surg Educ, Volume 73, page 245-9 (2016).
10. Z. Zhao, S. Voros, Y. Weng, F. Chang, and R. Li, Tracking-by-detection of surgical instruments in minimally invasive surgery via the convolutional neural network deep learning-based method, Computer Assisted Surgery, Volume 22, page 26-35 (2017).
11. M. Zhang et al., Multiple instruments motion trajectory tracking in optical surgical navigation, Optics Express, Volume 27, page 15827-15845, (2019).
12. D. Berry, Percutaneous aortic valve replacement: an important advance in cardiology, Eur Heart J., Volume 30, page 2167-2169 (2009).

13. S. Kobayashi et al., Assessment of surgical skills by using surgical navigation in robot-assisted partial nephrectomy, International Journal of Computer Assisted Radiology and Surgery, 2019/05/22 2019, doi: 10.1007/s11548-019-01980-8.
14. P. L. Docquier, L. Paul, and K. TranDuy, Surgical navigation in paediatric orthopaedics, EFORT Open Reviews, Volume 1, page 152-159 (2016).
15.
16. H. Tadayyon, A. Lasso, A. Kaushal, P. Guion, G. Fichtinger, Target motion tracking in MRI-guided transrectal robotic prostate biopsy, IEEE Trans. Biomed. Eng., Volume 58, page 3135-42 (2011).
17. E. Ozkan, C. Tanner, M. Kastelic, O. Mattausch, M. Makhinya, and O. Goksel, Robust motion tracking in liver from 2D ultrasound images using supporters, Int J Comput Assist Radiol Surg, Volume 12, page 941-950 (2017).
18. T. J. Liu, A. T. Ko, Y. B. Tang, H. S. Lai, H. F. Chien, and T. M. Hsieh, Clinical application of different surgical navigation systems in complex craniomaxillofacial surgery: The use of multisurface 3-dimensional images and a 2-plane reference system, Ann Plast Surg, Volume 76, page 411-9 (2016).
19. S. Engelhardt et al., Accuracy evaluation of a mitral valve surgery assistance system based on optical tracking, Int J Comput Assist Radiol Surg, Volume 11, page 1891-904 (2016).
20. R. Niehaus et al., Experience of total knee arthroplasty using a novel navigation system within the surgical field, Knee, Volume 24, page 518-524 (2017).
21. B. G. Kim and D. J. Park, Unsupervised video object segmentation and tracking based on new edge features, Pattern Recognition Letters, Volume 25, page 1731-1742 (2004).
22. B. N. Subudhi, P. K. Nanda, and A. Ghosh, A change information based fast algorithm for video object detection and tracking, IEEE Trans. on Cir. and Sys. for Video Tech., Volume 21, page 993-1004 (2011).
23. S. Duffner and C. Garcia, Fast pixel wise adaptive visual tracking of non rigid objects, IEEE Transactions on Image Processing, Volume 26, page 2368-2380 (2017).
24. J. Li, X. Zhou, S. Chan, and S. Chen, Robust object tracking via large margin and scale adaptive correlation filter, IEEE Access, Volume 6, page 12642-12655 (2017).
25. T. Mahalingam, M. Subramoniam, A robust single and multiple moving object detection, tracking and classification, Applied Computing and Informatics, Elsevier (2018).
26. T. Zhang, S. Liu, C. Xu, B. Liu, and M. Yang, Correlation particle filter for visual tracking, IEEE Transactions on Image Processing, Volume 27, page 2676-2687 (2018).
27. G. Yang yang, H. Dong-jian, and L. Cong, Target tracking and 3D trajectory acquisition of cabbage butterfly based on the KCF-BS algorithm, Scientific Reports, Nature, Volume 8, page 9622 (2018).
28. B. Du, Y. Sun, S. Cai, C. Wu, and Q. Du, Object tracking in satellite videos by fusing the kernel correlation filter and the three-frame-difference algorithm, IEEE Transactions on Image Processing, Volume 15, page 168-1821 (2018).
29. J. Ning, L. Zhang, D. Zhang, and W. Yu, Joint registration and active contour segmentation for object tracking, IEEE Transactions on Circuits and Systems for Video Technology, Volume 23, page 1589-1597 (2013).
30. G. Liu, S. Liu, K. Muhammad, A. Sangaiah, and F. Doctor, Object tracking in vary lighting conditions for fog based intelligent surveillance of public spaces, IEEE Access, Volume 6, page 29283-29296 (2018).
31. S. Liu and Y. Feng, Real-time fast moving object tracking in severely degraded videos captured by unmanned aerial vehicle, Int. Journal of Adv. Rob. Sys., SAGE, Volume 11, page 1-10 (2018).
32. D. Du, L. Wen, H. Qi, Q. Huang, Q. Tian, and S. Lyu, Iterative graph seeking for object tracking, IEEE Transactions on Image Processing, Volume 27, page 1809-1821 (2018).
33. Radau P., Lu Y., Connelly K., Paul G., Dick A. J., and Wright G. A. Evaluation framework for algorithms segmenting short axis cardiac MRI, The MIDAS Journal - Cardiac MR Left Ventricle Segmentation Challenge (2008).
34. S. Dakua, Julien Abinahed, and Abdulla Al-Ansari, A PCA based approach for brain aneurysm segmentation, Journal of Multi Dimensional Systems and Signal Processing, Springer, Volume 29, page 257-277 (2018).
35. Rallabandi V. and Roy P., MRI enhancement using stochastic resonance in Fourier domain, Magnetic Resonance Imaging, Volume 28, page 1361-1373 (2010).
36. J. vom Scheidt and T. C. Gard, Introduction to Stochastic Differential Equations, New York-Basel, Marcel Dekker Inc., XI, 234 pp., $ 78.-. ISBN 0-8247-7776-X (Pure and Applied Mathematics 114).
37. S. J. Yao, Y. H. Song, L. Z. Zhang, and X. Y. Cheng, MODWT and networks for short-term electrical load forecasting, Energy conversion and Management, Volume 41, page 1975-1988 (2000).
38. D. Comaniciu, V. Ramesh, and P. Meer, Kernel-Based Object Tracking, IEEE Trans. Pattern Anal. Machine Intell., Volume 25, page 564-577 (2003).
39. E. Vezzetti and F. Marcolin, Similarity Measures for Face Recognition, Bentham Books, ISBN: 978-1-68108-045-1 (2015).
40. C. E. Erdem, B. Sankur, and A. M. Tekalp, Performance measures for video object segmentation and tracking, IEEE Transactions on Image Processing, Volume 13, page 937-951 (2004).
41. Kristan Matej, Jiri Matas, Ales Leonardis, Michael Felsberg, Luka Cehovin, The visual object tracking VOT 2015 challenge results, in IEEE International Conference on Computer Vision Workshop, page 564-586 (2015).

42. Rallabandi V. P. and Roy P. K., MRI enhancement using stochastic resonance in Fourier domain, Magnetic Resonance Imaging, Elsevier, Volume 28, page 1361-1373 (2010).
43. Wang, Z., Sheikh, H. R., and Bovik, A. C., No-reference perceptual quality assessment of jpeg compressed images, in Proc. IEEE Int. Conf. Image Processing, page 477-480 (2002).
44. N. Paragios and R. Deriche, Geodesic active contours and level sets for the detection and tracking of moving objects, IEEE Trans. Pattern Anal. Machine Intell., Volume 22, page 266-280 (2000).
45. G. H. Rosenfield and K. Fitzpatrick Lins, A coefficient of agreement as a measure of thematic classification accuracy, Photogramm. Eng. Remote Sens., Volume 52, page 223-227 (1986).
46. Feng Shi, Qi Yang, Xiuhai Guo, Touseef Qureshi, Zixiao Tian, Huijuan Miao, Damini Dey, Debiao Li, and Zhaoyang Fan, Vessel Wall Segmentation Using Convolutional Neural Networks, IEEE Trans. on Biomedical Engineering. doi: 10.1109/TBME.2019.2896972
47. http://www.votchallenge.net/vot2015/dataset.html